Pergamon

0042-6989(95)00044-5

Vision Res. Vol. 35, No. 21, pp. 3003–3013, 1995

0042-6989/95 $9.50 + 0.00

Fast Perceptual Learning in Hyperacuity

M. FAHLE,* S. EDELMAN,† T. POGGIO‡

Received 31 January 1994; in revised form 23 December 1994

We investigated fast improvement of visual performance in several hyperacuity tasks such as vernier acuity and stereoscopic depth perception in almost 100 observers. Results indicate that the fast phase of perceptual learning, occurring within less than 1 hr of training, is specific for the visual field position and for the particular hyperacuity task, but is only partly specific for the eye trained and for the offset tested. Learning occurs without feedback. We conjecture that the site of learning may be quite early in the visual pathway.

Cortical plasticity Vernier acuity Monocular specificity Visual field position Stereoscopic acuity Motion detection

INTRODUCTION

Recent psychophysical evidence suggests that human adults are able to significantly improve performance in perceptual tasks. Improvement occurs within a training time of a few hundred presentations (Fiorentini & Berardi, 1980, 1981; Karni & Sagi, 1991; Shiu & Pashler, 1992; Poggio, Fahle & Edelman, 1992; Vaina, personal communication; Weiss, Edelman & Fahle, 1993). Learning is specific in a number of respects. For instance, the improvement of thresholds caused by perceptual learning disappears after rotation of the vernier stimulus by 90 deg (Poggio, Fahle & Edelman, 1991; Poggio *et al.*, 1992; Ahissar & Hochstein, 1993) and learning seems to be specific for visual field position (Ramachandran & Braddick, 1973; Nazir & O'Regan, 1990; O'Toole & Kersten, 1992). On the other hand, more cognitive aspects such as global pattern structure, attention and motivation will also influence performance and the speed of learning (e.g. Ahissar & Hochstein, 1993; Weiss *et al.*, 1993; Rentschler, Jüttner & Caelli, 1994; Herzog & Fahle, 1994).

In trying to clarify the mechanisms underlying perceptual learning, in this paper we investigate the specificity of perceptual learning with respect to visual field position, offset size and the eye tested. Since learning is specific for orientation, one may suspect that it may also be specific for visual field position, for the offset size used during training, and perhaps even for the eye that learned under monocular conditions. Another important question discussed in this paper is the role of feedback. In a previous investigation (Fahle & Edelman, 1993; see also Weiss *et al.*, 1993), we found perceptual learning even without auditory feedback, well in line with earlier reports in the literature (McKee & Westheimer, 1978). It remained unclear, however, whether learning without feedback is significantly slower than with feedback, and whether learning is possible even with presentation of sub-threshold stimuli only. The third goal of this paper was to find out whether the orientation specificity of vernier learning would also hold true for other domains such as stereoscopic depth perception.

MATERIALS AND METHODS

Stimuli, usually line or three-dot vernier targets, appeared on an oscilloscope screen (Hewlett Packard 1333; P31 phosphor) under computer control. Targets were 10 min arc long and 2 min arc wide, intensified every 16 msec to a mean luminance of 240 cd/m^2 on a background of 2 cd/m^2; i.e. contrast was $C_w = L_{max}/L_{min} = 120$. Stimuli appeared for 150 msec each, and observers had to indicate the direction of vernier offset by pressing the appropriate one of two push-buttons. The next stimulus followed 500 msec after the observer's response. The computer supplied auditory error-feedback except in Expt 4.

We started each experiment for each observer by measuring a threshold based on a block of 80 stimulus presentations, using a method of adaptive Probit estimation (PEST) (Taylor & Creelman, 1967). This threshold indicated the individual level of performance of each subject and was typically around 15 sec arc for 75% correct responses, with the extremes ranging between 7.5 and 25 sec arc. In the following blocks of the experiment, percentages of correct responses for the threshold displacement obtained in the first block were collected, with block size always being 80 presentations. This is to say that we used percentages of correct

*Section of Visual Science, Department of Neuroophthalmology, Waldhörnlestr. 22, D72076 Tübingen, Germany [*Email* manfred.fahle@commlink.uni-tuebingen.de].

†Department of Applied Mathematics & Computer Science, The Weizmann Institute, 76100 Rehovot, Israel.

‡Artificial Intelligence Laboratory, Massachusetts Institute of Technology, 555 Technology Square, Cambridge, MA 02139, U.S.A.

responses, rather than thresholds, in most of the experiments to assess the improvement in performance due to learning. Thresholds in earlier experiments had proven to be too variable in inexperienced observers to be used for this purpose, though we give an example of how thresholds improve in the experiment on stereoscopic depth perception.

Observation distance was 2.5 m, and positional accuracy on the screen better than 1 sec arc at this distance. Observers were paid volunteers, usually undergraduate students of Tübingen university who gave informed consent to participate in this study. They were naive as to the exact aim of the experiments. Testing lasted 2 hr in most of the experiments, divided to two sessions, normally on consecutive days. The change in stimulus or viewing conditions (e.g. right eye–left eye) always occurred after the first block of the second session, in order to separate possible effects of forgetting between the sessions from stimulus-specific effects. At the end of the second session, the first stimulus condition was retested once. Altogether, almost 100 observers participated in the experiments. Their visual acuity was normal or corrected-to-normal, and each observer participated in only one experiment.

To evaluate the results, we have performed an analysis of covariance (ANCOVA), using the Statistical Analysis Software (SAS) procedure General Linear Models (GLM), with a homogeneity-of-slopes model. The percentages of correct responses resp. thresholds in each block were first transformed logarithmically to accommodate the asymptotic improvement with time.

Simulations

We originally used a HyperBF network to test the computational plausibility of learning a task such as vernier acuity from signals provided by cells early in the visual pathway, such as photoreceptor signals (Poggio *et al.*, 1992).

The operation of the simplest version of this model, sometimes called radial basis function (RBF) network, can be understood intuitively by employing the notion of similarity. Consider a system confronted with a novel stimulus, drawn from a space of possibilities some of which are already familiar to the system (i.e. the appropriate responses for these stimuli are known). The required response can then be formed by combining the responses to familiar stimuli that are similar to the present one. To make this notion more precise, one must decide how to represent the stimuli, define similarity between stimulus representations, and specify the manner of combining responses to familiar stimuli. In the RBF model we used, a vernier is represented by the pattern of activities it evokes in a set of receptive fields onto which it is projected, similarity is defined in terms of distance between vectors of activities of the receptive fields, and the new response is computed as a linear superposition of responses to the familiar stimuli. The familiar stimulus–response pairs are stored in the system and are considered the basis functions or the templates generating the response to a new stimulus. In principle, a system implementing the RBF model can start from a *tabula rasa* situation and acquire the necessary basis functions by storing the incoming examples during a learning phase.

Formally the output of the RBF network is $y = C\mathbf{g}(\mathbf{x})$, where C is the weight matrix and $\mathbf{g}$ is a vector where each component is the output of one of a set of hidden units, that embody the basis functions. The input vector $\mathbf{x}$ consists, in our original model, of the signals from photoreceptors [Poggio *et al.* (1992) mentioned, however, as more biologically plausible a set of inputs similar to the signal of cortical cells both oriented and circularly symmetric]. Each unit performs an operation that could be described as "blurred" template matching by measuring the similarity of the novel input with the template which represents the optimal stimulus t to which the unit is tuned. The activity of the unit depends then on this similarity through (typically) a Gaussian function $G(\|\mathbf{x} - t\|)$. At the output of the network the activities of the various units are combined with appropriate weights, determined during the learning stage. Thus classification of a novel vernier stimulus is achieved by comparing the actual stimulus with the stored templates (which are identical with the examples in the RBF network) and by combining the results of these comparisons (Poggio *et al.*, 1992).

This is the simplest in a class of models called hyper basis function (HyperBF) networks (Poggio & Girosi, 1990; Girosi, Jones & Poggio, 1993). In the more general HyperBF scheme the number of units, i.e. templates, used during recognition may be less than the number of training views; the centers are also changed during learning and are in general different from any of the examples after the training phase; furthermore the appropriate similarity metric may be found automatically during learning. Even in a HyperBF network, however, the centers will be very similar to the examples whenever the training phase is as short as in all the simulations of this paper.

RESULTS

Experiment 1. Interocular transfer

Here, only one eye saw the stimuli. Six observers started with the left eye, six started with the right eye. After the first block of the second session, a short pause (3 min) was made and testing continued with the other eye. An opaque occluder, as used for perimetry and, which did not exert any pressure on the occluded eye, covered one eye. In control experiments with additional observers, transition between eyes occurred before the start of the second session, with very similar results. As it is obvious from Fig. 1(a), learning improved performance significantly even under monocular conditions, and learning transferred only partly to the partner eye, i.e. fast perceptual learning in vernier acuity is at least in part specific for the eye through which the learning occurred. Figure 1(b) illustrates the variability of the effect across different subjects. This topic and the caution needed in interpreting the data is discussed later.

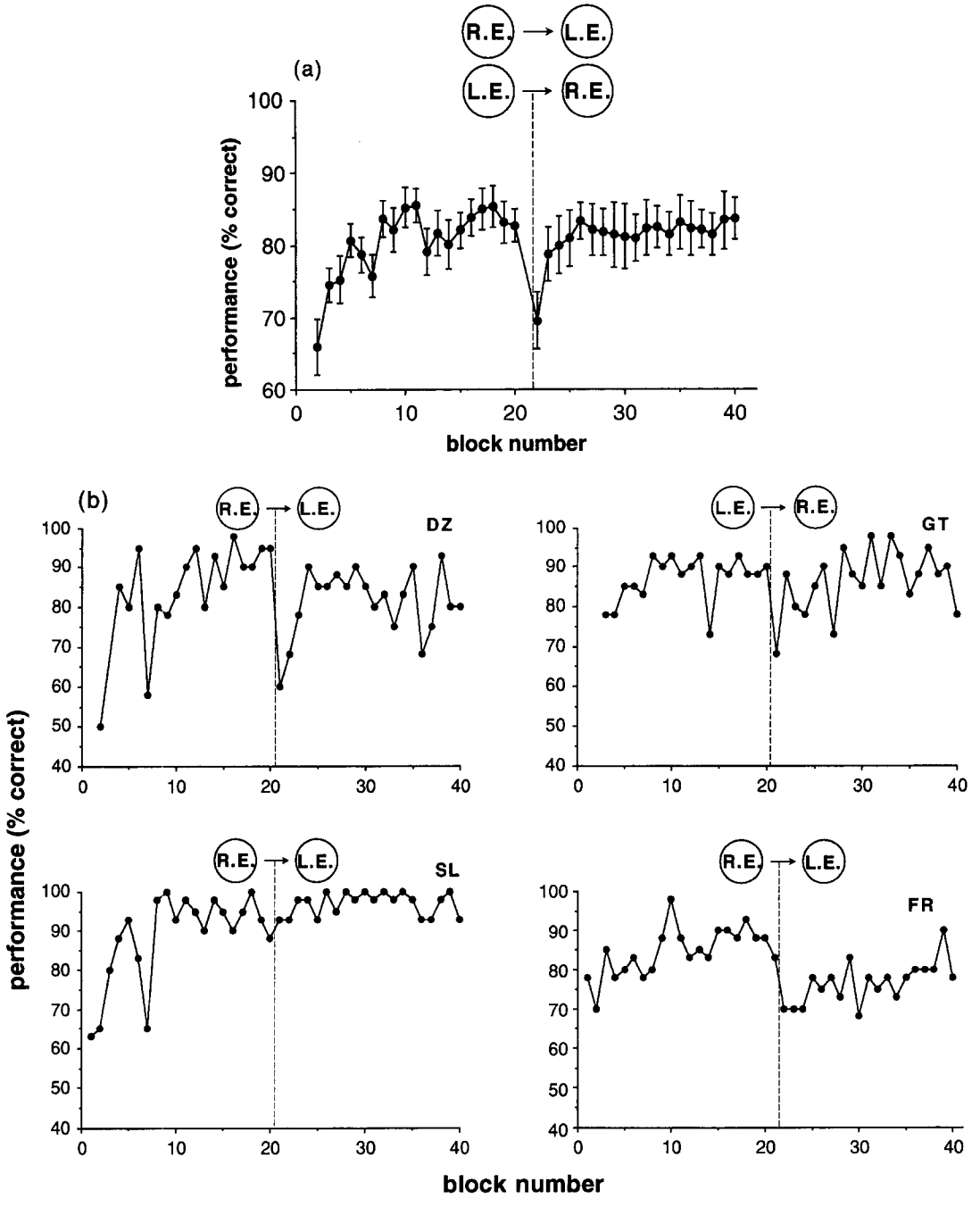

FIGURE 1. (a) Vernier performance as a function of monocular learning. Means and SEs of 12 observers. Six observers started with the left eye patched while the remaining six observers initially had the right eye patched. After 1 hr the eye patch was moved to the other eye. The average data seem to show that learning did not completely transfer from one eye to the partner eye. The effect of eye, as evaluated by an ANCOVA (see Materials and Methods) was not significant [$F(1,464) = 2.25$, $P = 0.134$] and neither was the interaction of block and eye [$F(1,464) = 1.76$, $P = 0.186$]. Only the effect of block reached significance [$F(1,464) = 23.3$, $P < 0.0001$]. This is to say that the transfer of learning between the eyes failed to reach significance. (b) Individual results of four of the 12 observers that are pooled together in (a). The first one (DZ) shows a clear absence of transfer, the second (SL) a clear transfer and the third and forth (GT and FR) are typical for the remaining 10 of the subjects, each one of which does not show a clear trend.

Experiment 2. Visual field position

A pilot study in which vernier targets appeared at a peripheral visual field position and then at another position gave inhomogeneous results, with sometimes complete transfer of learning from one position to another even when orientation changed by 90 deg at the same time as visual field position changed. We concluded that two mechanisms might be active under these conditions: one that was stimulus specific and another one that might be concerned with moving attention from the center to the periphery of the visual field. We therefore decided to measure performance sequentially at eight visual field positions. The sequence of visual field positions, all at an eccentricity of 10 deg, was counterbalanced among the eight observers, who had previously trained another perceptual task in the periphery. A video camera and video recorder monitored eye position for

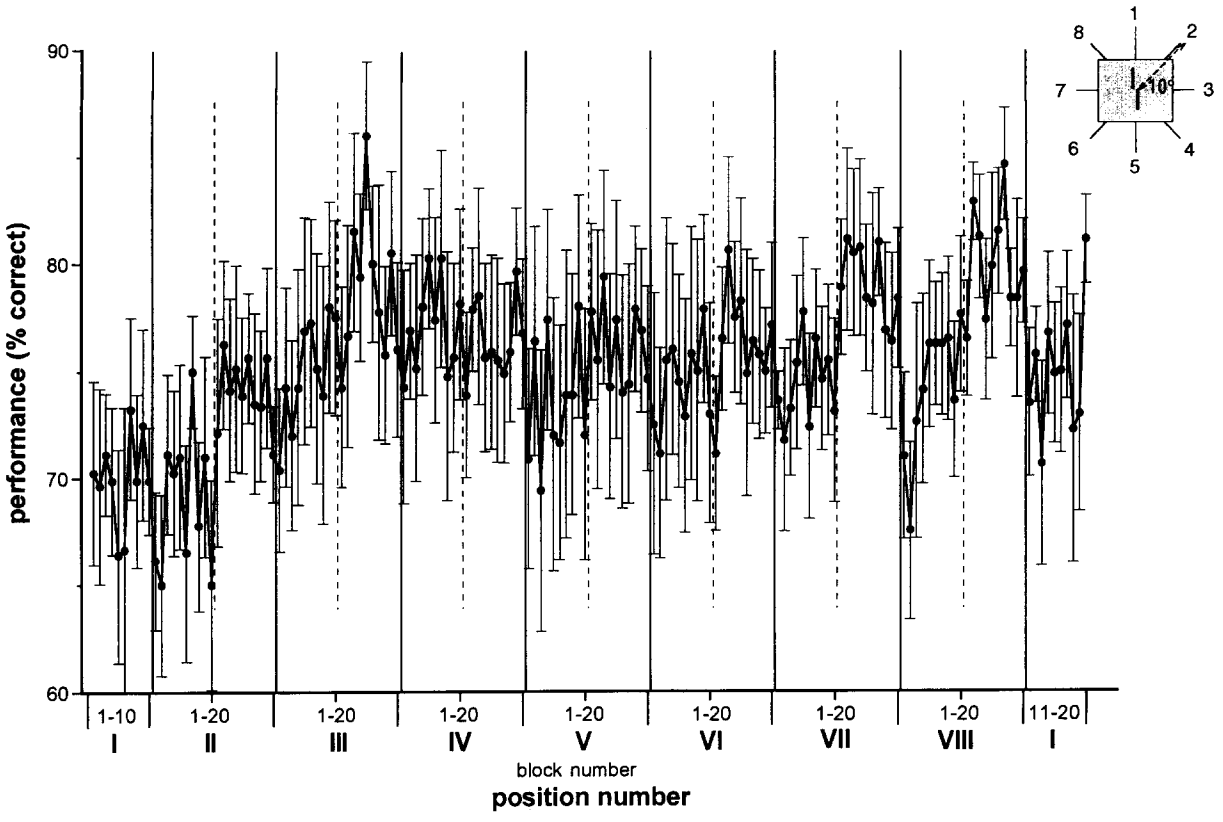

FIGURE 2. Performance in vernier detection as a function of binocular learning in the periphery of the visual field. The vernier target was presented 80 times at each of the visual field positions at 10 deg eccentricity (Nos 1–8). Visual field position changed after each hour of training. Inset shows the visual field positions tested. Observers fixated on one of the positions indicated by the Nos 1–8. Sequence of positions was counterbalanced between the eight observers that participated in the experiment, i.e. at each period of learning (I–VIII), each observer fixated a different position, and each observer learned at each position only once (except at his/her first position, I). The graph shows that mean results of the eight observers improved during most of the 1 hr intervals, but decreased at most of the transitions between positions (vertical lines). The breaks between sessions however (interrupted vertical lines) did not significantly influence the results. The effect of position was significant [$F(7,1264) = 2.53$, $P = 0.0136$], i.e. there was an over-all improvement with time. The effect of block was highly significant [$F(1,1264) = 28.2$, $P = 0.0001$], i.e. observers improved at each position, but the interaction between block and position failed to reach significance ($P = 0.35$), indicating that the extent of learning was similar for all positions tested.

fixation control. Means and SDs of all observers are shown in Fig. 2, while Fig. 3 shows the results of all observers, collapsed over subsequent transitions. Learning of vernier acuity appears to be specific for visual field position. The interval between subsequent sessions of usually 24 hr, however, did not influence the results significantly.

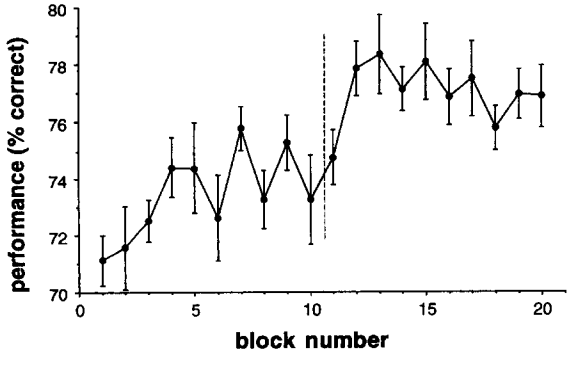

FIGURE 3. Vernier detection in the periphery of the visual field. Results of all observers shown in Fig. 2 have been collapsed over subsequent changes of visual field position (i.e. over positions 1–8). The dashed line indicates again the break between subsequent sessions, i.e. usually subsequent days. In this graph the improvement of performance is more clearly seen than in the previous graph.

Experiment 3. Stimulus range

The HyperBF model with photoreceptor inputs (see Methods) would predict that learning of vernier acuity is at least partly specific for the offset size presented during learning. To test this prediction, we performed two experiments. In the first experiment, one group of six observers had to discriminate between verniers offset to the right vs to the left by 10 sec arc, and after 1 hr, offset was enlarged to be 30 sec arc [Fig. 4(a)]. A second group of six observers started with stimuli offset by 30 sec arc and after 1 hr, offset was set to 10 sec arc [Fig. 4(b)]. Thus for each block of Fig. 4(c), there was an equal number of 10 and 30 sec arc stimuli. The combined results of both groups show that there is an almost complete transfer of learning from one offset size to another: there is only a small dip in performance following the transition between offset sizes. As expected, simulations of the HyperBF model as shown in Fig. 4(d) show a much sharper dip in performance that seems at odds with our experimental results.

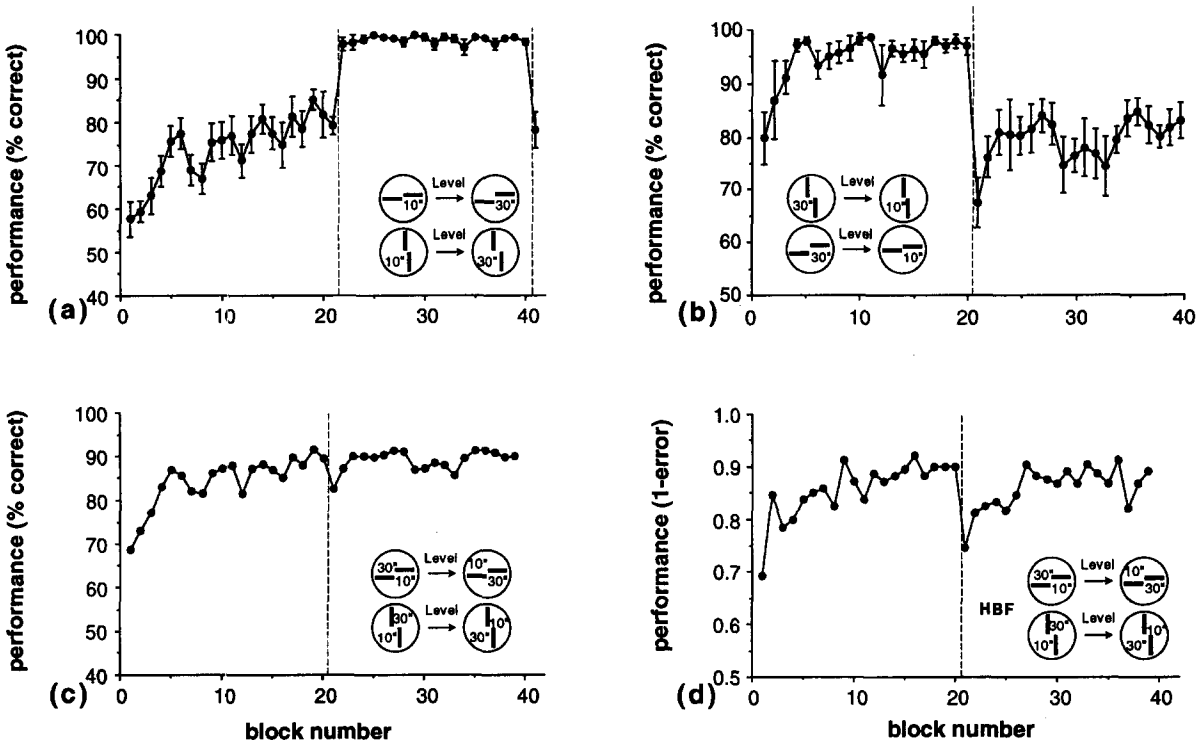

FIGURE 4. Vernier performance for stimulus offsets of 10 and 30 sec arc. One group of 6 observers started with an offset of 10 sec arc and switched to 30 sec arc after 1 hr of training (a), while the sequence was reversed for the second group of observers (b). Means of all observers are shown in (c). (d) The simulation of the experiment using a HyperBF network. The input to the network was an array of "photoreceptor" activities (the identification of the input units with photoreceptors here does not imply an attempt to model the neurophysiological aspects of hyperacuity perception). There were eight receptors, positioned randomly on a loose 4×2 grid. Each of the receptors calculated its response by integrating the input over a Gaussian-shaped region of the "retina". The eight-component vector of receptor outputs constituted the input to the RBF module, which was trained to produce an output of $+1$ for one sense of the input vernier displacement, and -1 for the other. During testing the performance was measured by counting the proportion of trials in which the sign of the module's output was consistent with the sense of the vernier displacement. The effect of levels was highly significant both in (a) and (b) [$F_a(1,235) = 335$, $P = 0.0001$; $F_b(1,236) = 33.6$, $P = 0.0001$], as was the effect of block [$F_a(1,235) = 36.5$, $P < 0.0001$; $F_b(1,236) = 14.4$, $P = 0.0002$]. The interaction between block and level was significant for (a) [$F(1,235) = 37.0$, $P < 0.0001$] but not significant for (b) [$F(1,236) = 0.54$, $P = 0.463$]. Here, the difference was obviously caused by the ceiling effect in (a).

In the second part of the experiment, six observers started with an offset size of 30 sec arc that decreased to 20 sec arc after 30 min of training, and to 10 sec arc after 60 min of training. Therefore, this group had to detect offsets of 10 sec arc after 1 hr of training, exactly as the second group of observers in the preceding experiment [Fig. 4(b)]. The only difference was that training in the previous experiment had been exclusively with 30 sec arc offsets, while it was equally divided to 30 and 20 sec arc in this experiment. The results of the observers in Fig. 4(b) are somewhat lower at the 21st block, i.e. immediately after the transition to the smallest offset, than the results for the corresponding block in Fig. 5. This result indicates a limited capability of observers to generalize between offset ranges. The difference between groups of observers, however, is not significant for the 21st block.

Experiment 4. Influence of feedback

Previous investigations have shown that learning can occur even without feedback (McKee & Westheimer, 1978; Fahle & Edelman, 1993; Shiu & Pashler, 1992; Karni & Sagi, 1991). To test whether this is also true for offsets that initially range below threshold, and to assess the speed of learning in the absence of feedback, we presented vernier stimuli with a constant offset of 10 sec arc, oriented either vertically or horizontally. Six observers received auditory feedback regarding the correctness of their responses, another six observers did not receive feedback. Figure 6 shows the results. Learning occurred in the absence of feedback [Fig. 6(a)], even if offsets were initially below threshold. The slope of log percentage correct in the feedback condition was 0.0017, compared to 0.0011 in the no-feedback condition. The data averaged over the six subjects in each condition showed a similar pattern. A regression of log percentage correct on block for the averaged data yielded slopes of 0.0016 (different from 0 at $P = 0.0001$) and 0.0011 (different from 0 at $P = 0.0017$) in the feedback and no-feedback conditions respectively. These slopes differ not significantly from each other ($P = 0.156$).

Experiment 5. Transfer between line and point vernier stimuli

In principle, learning might be specific for the stimulus presented, or might lead to a better spatial resolution, at least for the orientation learned. To discriminate between the two possibilities, we investigated whether

learning would transfer from one hyperacuity stimulus, namely standard line verniers, to a three-point vernier stimulus of identical orientation and size, and vice versa. The inset of Fig. 7 indicates the shape of the stimuli. The results of 12 observers indicate that there is hardly any transfer of learning between the two hyperacuity tasks. Figure 7(d) shows the result of simulations of a HyperBF network on the same task. In this case the simulations are more consistent with the experimental data: there is a dip in performance, even though it is smaller than in the observer's results.

In a related task, three observers had to indicate whether a vernier stimulus was offset to the left or was straight. After 1 hr of training, the task changed to the discrimination between a straight vernier and a vernier offset to the right. For another three observers, the sequence of tasks was reversed. As shown in Fig. 8(a), there was partial transfer of learning from offsets to the right to offsets to the left. Simulations with a HyperBF network give quite different results, since they show a very significant drop in performance [Fig. 8(b)].

Experiment 6. Orientation dependence in other hyperacuity tasks

We have reported previously that learning in vernier acuity is specific for orientation both for the fast phase of learning (Poggio *et al.*, 1992) and the slow phase of learning (Fahle & Edelman, 1993). The same is true for learning in a jump displacement task (Fahle & Skrandies, 1995; Fahle, 1994). Stereoscopic thresholds were measured in six observers to assess whether the orientation specificity of learning holds true also for three-dimensional space perception.

Stimuli appeared on two x/y monitors, equipped with polarizing filters matched to filters in front of the observers' eyes. A semi-transparent mirror superimposed the images of both observers (cf. Fahle & Westheimer, 1988). An adaptive method of sequential testing determined thresholds (PEST, see above). The observer's task was to detect which of two points was closer to the observer and which one was farther away. Five observers started with the points being separated horizontally by 10 min arc, and switched to a vertical arrangement of the points after 1 hr of training, while the other five observers started with a vertical arrangement of the points. Thresholds improved significantly as a function of learning, but no transfer of learning occurred from one stimulus orientation to a stimulus rotated by 90 deg (Fig. 9).

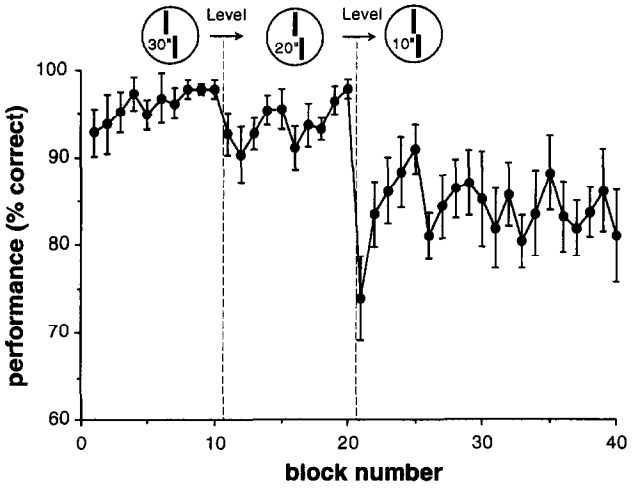

FIGURE 5. Vernier performance for different offsets, as in Fig. 4(b). Here, however, an offset size of 20 sec arc was tested between the 30 and 10 sec arc offsets. After 1 hr of training conditions were identical for the observers whose results are shown here and the observers of Fig. 4(b). Results for block No. 21, however, are not significantly better for the observers of Fig. 5 (unpaired t-test, $P = 0.19$). The effect of level is significant [$F(2,234) = 7.58$, $P = 0.0006$] as is the effect of block [$F(1,234) = 4.07$, $P = 0.045$], while the interaction between block and level is not significant.

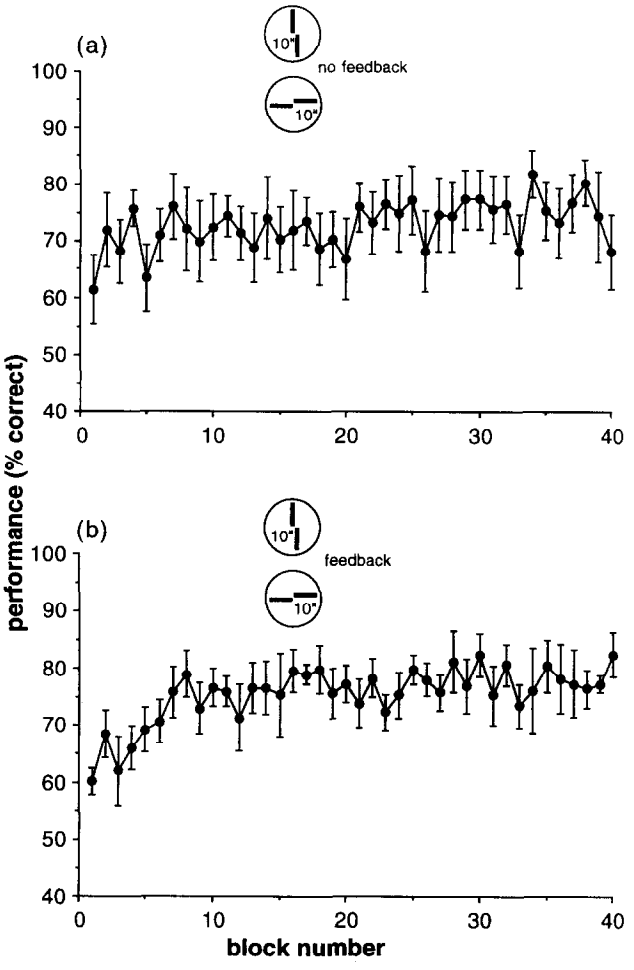

FIGURE 6. Vernier performance for subthreshold offsets (10 sec arc) as a function of learning (a) in the absence of auditory feedback (six observers) and (b) in the presence of auditory feedback (six observers). Means and SEs of all observers of each group. The interaction between the effects of feedback and of block was highly significant [$F(2,476) = 11.2$, $P < 0.0001$], i.e. there was a highly significant difference between results of observers who received auditory feedback versus those who did not receive error feedback.

DISCUSSION

Most of our experiments aimed to characterize the specificity of perceptual learning which is its most surprising feature and the one that suggests an early locus of learning in the visual pathway. Our results support our own preliminary reports and results by others in

similar perceptual tasks: fast learning of hyperacuity tasks does not fully transfer from an orientation to a different one, from one eye to the other, from one position in the visual field to another and from point verniers to line verniers or vice versa. The specificity for stimulus orientation is a strong argument against the suspicion that observers actually learn a strategy, be it for the general experimental set-up, be it for more stable fixation or accommodation. Caution is needed, however, in interpreting in detail the general conclusion of specificity of perceptual learning, mainly because the variability in the data is quite high.

Inter-subject variability

Our results show a relatively high degree of noise and of individual variability, in line with previous reports (Fahle & Edelman, 1993; Kumar & Glaser, 1993). Notice that because of the nature of the learning task it is impossible to average, as usual in psychophysical experiments, equivalent multiple trials in the same individual. The data we present are therefore average data across several individuals. Their interpretation requires special care, since there is individual variability in the properties of the learning effects of which we ignore the causes—though inter-individual differences in the speed of learning are not an uncommon phenomenon also in everyday life.

Consider, in particular, the experiment about interocular transfer, which is probably the one that most suffers from the problem of individual variability. The data, averaged across 12 individuals, seem to suggest that transfer is far from complete. On the other hand, inspection of individual data sets suggests that there are only one or two observers that clearly show lack of transfer, whereas another one shows full transfer and each of the remaining nine or 10 plots is too noisy to support by itself either claim. Figure 1(b) contains four examples from the 12 time–courses we measured in as many subjects to illustrate the point that extrapolating from the average data to any specific individual may be quite dangerous. Additional data, however, on long-term learning, clearly show that at least the slow phase of vernier learning is eye specific (Fahle, 1994). These data, together with results by Sagi and Karni (1993) indicate that the slow phase of learning is specific for the eye that learned while the fast phase is far less specific and differs between observers.

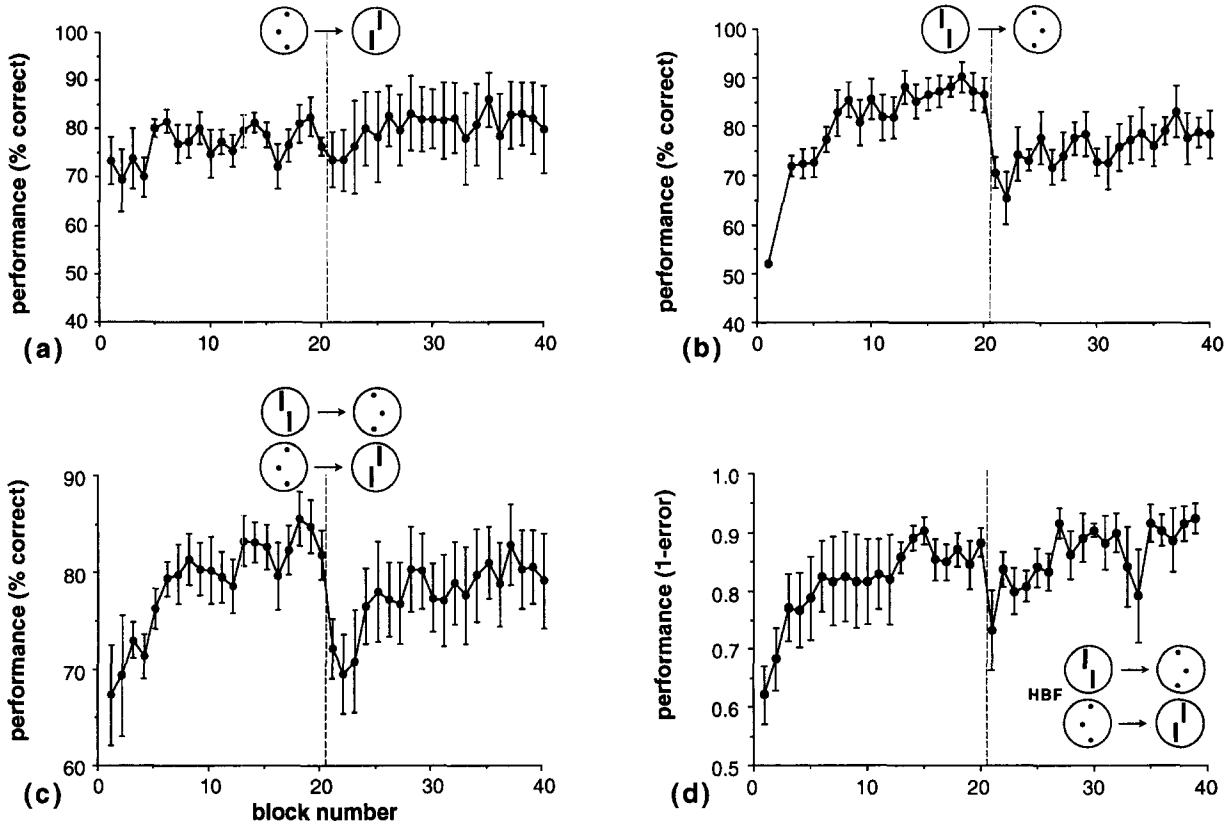

FIGURE 7. Vernier performance for standard line verniers and for three-dot verniers of identical orientation and over-all size. Six observers started with three-dot verniers. After the first block of the second hour of testing, these observers continued with vernier stimuli (a), while the experiment had the reverse temporal order for six other observers (b). Mean results clearly demonstrate that there is no transfer between the tasks (c). (d) The simulation with a HyperBF network. The effect of task was significant in (a) and (c) [$F_a(1,227) = 0.84$, $P = 0.04$; $F_b(1,468) = 7.27$, $P = 0.073$] while it was not significant in (b) ($P = 0.36$). The effect of block was significant in (a)–(c) [$F_a(1,227) = 4.04$, $P = 0.046$; $F_b(1,222) = 41.0$, $P < 0.0001$; $F_c(1,463) = 45.6$, $P < 0.0001$]. The interaction between block and task was significant only in (b) and (c) [$F_a(1,227) = 0.58$, $P = 0.31$; $F_b(1,222) = 8.1$, $P = 0.005$; $F_c(1,463) = 8.63$, $P = 0.0035$]. This is to say that there is a certain amount of transfer between the two tasks. The slope of a regression line through the first task is 1.13 ± 0.17 (%/block) ($P < 0.0001$), while the slope of a regression through the second task's results is 0.45 ± 0.16 ($P < 0.0064$).

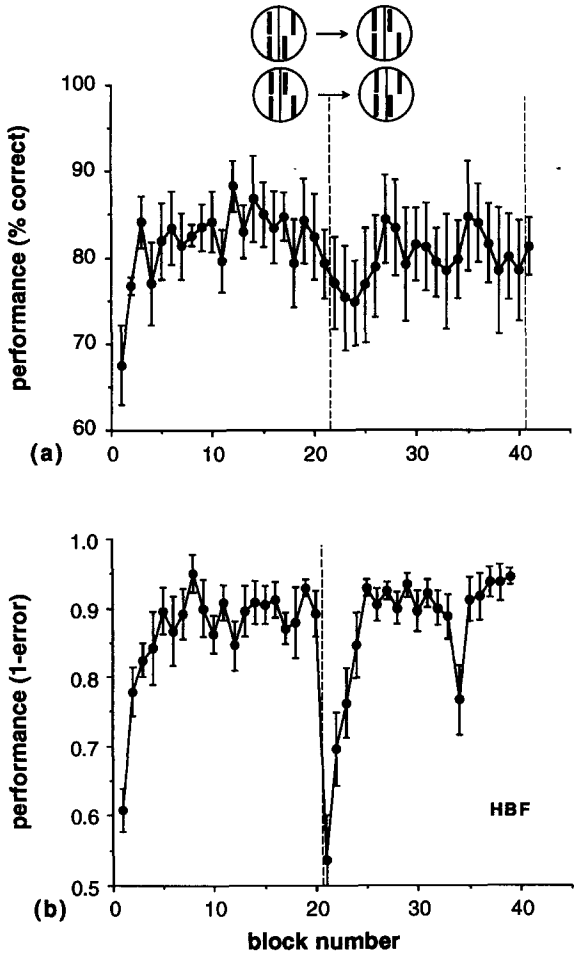

FIGURE 8. Vernier thresholds for the discrimination between a straight stimulus and a stimulus offset to the right (resp. left). After 1 hr of training the task changed to the discrimination between a straight vernier and a vernier offset to the left (resp. right). Individual results of six observers and their means (±SEs). (b) The result of simulations with a simple HyperBF network. The effect of block and the interaction between block and task are not significant in (a) ($P_b = 0.386$, $P_{b,t} = 0.328$), while the effect of task is significant [$F(1,336) = 6.8$, $P = 0.0095$], i.e. there is some transfer between the tasks. The slope of a regression line through the results of the first task is 0.422 ± 0.1667 (%/block) ($P = 0.017$) while the slope of a linear regression through the results of the second task is 0.191 ± 0.167 (n.s.).

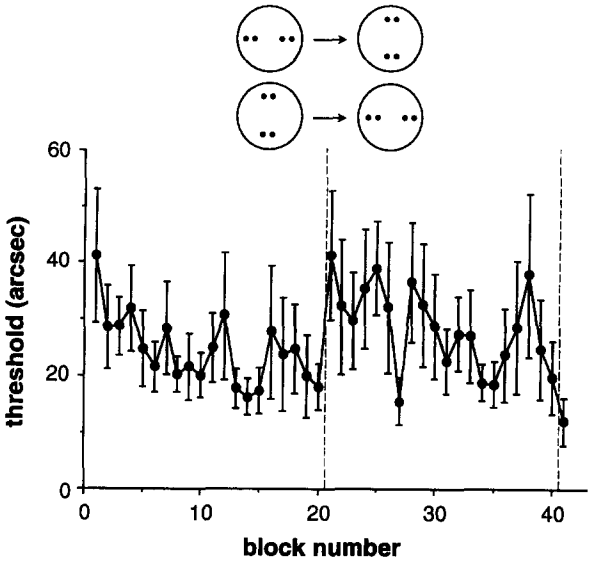

FIGURE 9. Stereoscopic thresholds as a function of learning. The first group of observers had to indicate whether the right or the left point was closer to them. After 1 hr of training, the points appeared with identical binocular disparities, but arranged vertically one above the other. The second group of observers underwent the reversed sequence of testing. There was no transfer of learning between the orientations. The effects of orientation and the interaction between block and orientation are not significant ($P_o = 0.41$, $P_{b,o} = 0.97$) while the effect of block is significant [$F(1,236) = 8.21$, $P = 0.0045$].

Models of perceptual learning

In this paper we attempt to characterize several properties of perceptual learning that are interesting in themselves and critical for the development of a model of the biological processes underlying perceptual learning. The model that we used originally (Poggio *et al.*, 1992) to motivate our experiments—a HyperBF network with photoreceptor-like inputs, see Fig. 10—was meant to provide a plausibility proof that perceptual tasks of the type described here can be indeed learned from a small set of examples. The argument was computational in nature: there is sufficient information in a few examples of the task to generalize to novel instances. The model also supported the argument that learning may be very specific, with little transfer to slightly different tasks, provided that the inputs used by the learning module were sufficiently low level.

Results of some of the experiments described in this paper are inconsistent with the predictions of the HyperBF network with photoreceptor-like inputs that we used originally (model 1). It is therefore interesting to

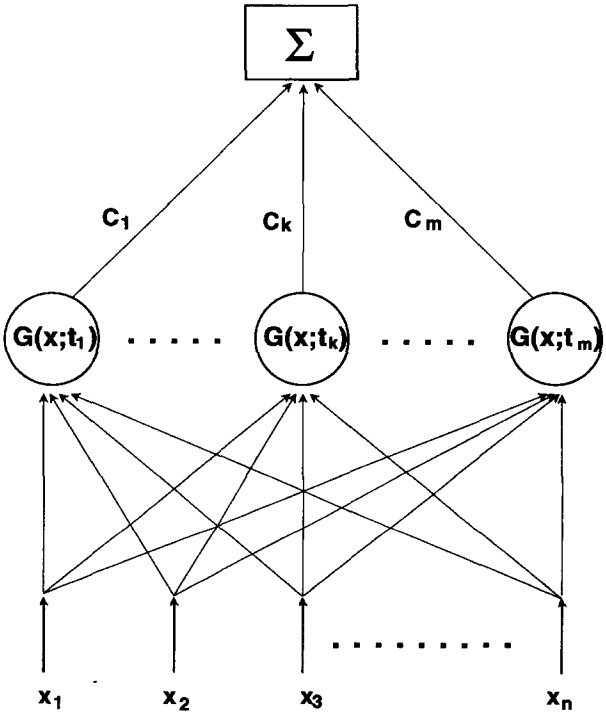

FIGURE 10. A network for interpolation with basis functions. The output is C $\mathbf{g}(\mathbf{x})$, where C is the weight matrix and $\mathbf{g}$ is a vector of hidden unit activities. Each hidden unit computes the similarity between the input and its preferred stimulus. In an RBF network, the activity of the hidden unit then depends on this similarity through a Gaussian function $G(\|\mathbf{x} - t\|^2)$.

consider alternative models such as a network that simply combines oriented receptive fields "looking" at the visual field with weights that are modified during learning (model 2). Weiss *et al.* (1993) originally proposed this model, by adding a learning component (the modifiable coefficients) to the "hardwired" model of Wilson (1986), which for the sake of the later discussion will be called model 3 (Wilson did not discuss the issue of learning of course and may correctly disagree with our hardwired interpretation of his model that we will carry on here for the sake of the discussion). Weiss *et al.* also demonstrated that their network (model 2, which is itself a HyperBF network with centers identical to oriented bars, see later) behaves in simulations in a way which is indistinguishable from a HyperBF model with oriented inputs and radial centers actually suggested (but not implemented) by Poggio *et al.* (1992).

Models and experiments

Three experiments showed evidence for some transfer of learning. Figure 8 suggests that subjects could generalize at least partly from the comparison straight-left to straight-right. Simulations with a HyperBF model with photoreceptor-like inputs are inconsistent with this experimental result.

Transfer from one range to another (see Figs 4 and 5) is less difficult to explain with a HyperBF model with photoreceptor-like inputs. Simulations show that when the range changed (increased or decreased) by a factor of 3, the dip in the performance is not as pronounced as in the previous experiment. Human data suggest that transfer is more difficult as the difference in range increases: there is more transfer from 20 to 10 sec arc than from 30 to 10 sec arc, though the difference is not significant. On the other hand there seems to be full transfer from 10 to 30 sec arc which seems intuitive enough but runs counter to the prediction of a simple memory-based model which may have learned examples of the 10 sec arc task but should find them too dissimilar to novel 30 sec arc stimuli to be useful for their classification.

The HyperBF model also exhibited a certain capacity for generalization across tasks: when the stimulus changed from line vernier to three dots, the dip in the performance was small, while it was more pronounced for the human subjects.

Reflections on models

The three models mentioned earlier are quite similar—so similar that they can all be regarded as specific instances of HyperBF networks. What then are their critical differences in the context of perceptual learning? As alluded to by Weiss *et al.* (1993) in their discussion, the key difference boils down to the old trade-off between *nature* and *nurture*. The hardwired model of Wilson is representative for the first extreme possibility; our model 1 for the latter. Model 2—the model of Weiss *et al.* (1993)—is a compromise between the two extreme hypotheses.

Our original experiments, on fast perceptual learning, demonstrated the inadequacy of a class of explanations such as model 3: hardwired networks are clearly inconsistent with data showing perceptual learning. The experiments of this paper suggest that explanations relying on learning mechanisms starting from a *tabula rasa* state—such as our original model 1—are also inadequate.

Model 1 starts from a *tabula rasa* situation creating new centers as the incoming examples require. It has therefore only the information contained in the set of examples. Models such as that of Weiss *et al.* (model 2), as well as models mixing 1 and 2 that we favor, have a repertoire of oriented receptive fields with different degrees of tuning, from coarse to sharp *before* the learning phase. Models of this type can be considered as HyperBF networks that start the learning phase with an existing *vocabulary* of centers, i.e. templates, tuned to a variety of optimal stimuli including bars of different orientations. The centers may or may not be modified during the actual learning phase; the coefficients c usually are (see Poggio & Girosi, 1989).

Thus the difference between the class of models that Weiss *et al.* (1993) propose and the original model we described (Poggio *et al.*, 1992) consists of whether the system can rely or not upon task-relevant information prior to being exposed to the set of examples. This prior information is embedded in a set of appropriate templates available *before* the training phase. The tuning of the associated cells may in turn depend on an earlier and much slower phase of learning at the level of the development of the individual or the species (a HyperBF network may "find" oriented templates as the optimal centers if provided with a sufficiently large set of examples). As Weiss *et al.* (1993) remark, model 2 is equivalent to the idea of innate perceptual mechanisms tuned by experience. Of course there may exist perceptual tasks and even hyperacuity tasks that are so unusual to render useless any prior information acquired by the organism prior to the task. In such a case—if it exists—we predict that the experimental results may be explained by a mixed 1 + 2 model synthesizing new centers in a stimulus-driven and task dependent way, in addition to an available vocabulary of existing tuned units. For the vernier and other hyperacuity tasks we have considered, however, the class of models strictly of the type of model 2 seems sufficient: they are consistent with the data of this paper and with the data available so far. They are also qualitatively consistent with the fact that adults are able to perform the vernier task for sufficiently large offsets even without any traning phase, supervised or unsupervised and that the effects of short-term learning are less pronounced in hyperacuity learning than in some other, more artificial learning tasks. These models are furthermore capable to better explain the results of experiments without feedback.

Experiments with and without feedback

Probably the most critical experiment for determining plausible classes of models is the experiment on the effect of feedback during learning. Figure 6 shows results that are consistent with the emerging agreement in the field

of perceptual learning: also in the case of hyperacuity tasks, perceptual learning seems to take place even in the absence of feedback and it may require relatively large datasets to obtain a significant difference between learning with vs without feedback (cf. Fahle & Edelman, 1993). The result implies that a large class of network models for learning cannot be used directly to develop a biologically plausible model of this phenomenon. In particular, HyperBasis function models, as well as Multilayer Perception networks, cannot be used directly since they both require a teacher signal during learning, i.e. feedback.

We had earlier suggested (Poggio *et al.*, 1992; see especially Weiss *et al.*, 1993) that networks of that type could still account for the data if used in a bootstrapping mode, in which very few initial examples correctly labeled could be sufficient to classify novel examples that are sufficiently similar to them. The experiments of Fig. 6 show that learning occurs even when the offsets are below thresholds at the beginning of the experiment. This seems at first glance to rule out a bootstrapping hypothesis (but not the learning method called EDL by Weiss *et al.*, or exposure-dependent learning; this rule relies on prior information about the task in that it keeps changes proportional to the existing weights). We should remember, however, that subjects still performed above chance even if below the psychophysical threshold. The experiments of Fig. 6 suggest in fact that the absence of feedback makes learning slower and may reduce its asymptotic performance. In a previous investigation on slow perceptual learning (Fahle & Edelman, 1993), slopes of regression lines through the data obtained with vs without feedback were significantly different from each other in one experiment but failed to reach significance in another experiment. We have now reanalyzed these data with an ANCOVA, using the SAS procedure GLM, with a homogeneity-of-slopes model specification. The results show significant differences between the feedback vs no feedback condition for both of the old data sets (Expts 1 and 2 vs Expt 3, $P = <0.05$; Expt 4 vs Expt 5, $P = <0.05$ for raw data). The dependence on feedback is thus consistent both with a bootstrapping hypothesis and the EDL rule. We remind that in *tabula rasa* HyperBF-like models learning takes place in two distinct ways: *unsupervised* learning is required to establish, create or tune the "centers" (unless they already all exist, e.g. as neurons tuned to many different optimal stimuli), whereas *supervised* learning determines the appropriate "synaptic" weights for the coefficients c (cf. Fig. 10). The first type of learning does not require feedback while the second does. Of course, if appropriate centers are already available prior to learning—e.g. as units tuned to oriented bars—bootstrapping may be possible even without any labeled examples.

There are, of course, other models that do not require feedback. We plan to explore such models in the context of a theory that will take into account not only the psychophysical data but also physiological and biophysical constraints. Here we remark only that at some level of abstraction some of these unsupervised models are equivalent to a network such the HyperBF-like network of Fig. 10 in which the coefficients c have already been set and the centers of the basis functions *adapt* their tuning—such as directional tuning of oriented cells—in an unsupervised fashion as the system is exposed to new (unlabeled) examples of the task.

The experiments summarized in Fig. 3 point to another aspect of perceptual learning that is relevant for any biological model: the time-course of learning. It seems that within one session learning reaches an asymptotic performance that can, however, improve rapidly the day after. This double-sigmoid behavior makes intuitive sense and may be connected with the (different) findings of Karni and Sagi (1991). Clearly, no period of rest is required to obtain improved performance in our task, and no improvement seems to occur during the period of rest. One might speculate that two opponent processes are evident in the graphs presented in this paper: learning improves performance, while fatigue tends to decrease performance during each session.

Our results show that perceptual learning is quite variable between observers, but that it is specific in most observers for the visual field position, for the orientation of the stimulus, for the eye that learned and for the exact task. Learning does not require feedback, even with subthreshold stimuli, and can be described by means of a modified HyperBF-based model that uses prior information about visual stimuli.

REFERENCES

Ahissar, M. & Hochstein, S. (1993). Attentional control of early perceptual learning. *Proceedings of the National Academy of Science, U.S.A., 90*, 5718–5722.

Fahle, M. (1994). Human pattern recognition: Parallel processing and perceptual learning. *Perception, 23*, 411–427.

Fahle, M. & Edelman, S. (1993). Long term learning in vernier acuity: Effects of stimulus orientation, range and of feedback. *Vision Research, 33*, 397–412.

Fahle, M. & Skrandies, W. (1995). An electrophysiological correlate of learning in motion perception. *German Journal of Ophthalmology*. In press.

Fahle, M. & Westheimer, G. (1988). Local and global factors in disparity detection of rows of points. *Vision Research, 28*, 171–178.

Fiorentini, A. & Berardi, N. (1980). Perceptual learning specific for orientation and spatial frequency. *Nature, 287*, 43–44.

Fiorentini, A. & Berardi, N. (1981). Learning in grating waveform discrimination: Specificity for orientation and spatial frequency. *Vision Research, 21*, 1149–1158.

Girosi, F., Jones, M. & Poggio, T. (1993). Priors, stabilizers and basis functions: From regularization to radial, tensor and additive splines. Massachusetts Institute of Technology, A.I. Memo No. 1430.

Herzog, M. & Fahle, M. (1994). Learning without attention? In Elsner, N. & Breer, H. (Eds), *Proceedings of the 22th Gottingen Neurobiology Conference* (p. 817). Stuttgart: Thieme.

Karni, A. & Sagi, D. (1991). Where practice makes perfect in texture discrimination: Evidence for primary visual cortex plasticity. *Proceedings of the National Academy of Science, U.S.A., 88*, 4966–4970.

Kumar, T. & Glaser, D. A. (1993). Initial performance, learning and observer variability for hyperacuity tasks. *Vision Research, 33*, 2287–2300.

McKee, S. P. & Westheimer, G. (1978). Improvement in vernier acuity with practice. *Perception & Psychophysics, 24*, 258–262.

Nazir, T. A. & O'Regan, J. K. (1990). Some results on translation invariance in the human visual system. *Spatial Vision, 5*, 81–100.

O'Toole, A. & Kersten, D. (1992). Learning to see random-dot stereograms. *Perception, 21*, 227–243.

Poggio, T. & Girosi, F. (1990). Regularization algorithms for learning that are equivalent to multilayer networks. *Science, 247*, 978–982.

Poggio, T., Fahle, M. & Edelman, S. (1991). Synthesis of visual modules from examples: learning hyperacuity. Artificial Intelligence Memo, 1271, MIT, Cambridge, Mass.

Poggio, T., Fahle, M. & Edelman, S. (1992). Fast perceptual learning in visual hyperacuity. *Science, 256*, 1018–1021.

Ramachandran, V. S. & Braddick, O. (1973). Orientation-specific learning in stereopsis. *Perception, 2*, 371–376.

Rentschler, I., Jüttner, M. & Caelli, T. (1994). Probabilistic analysis of human supervised learning and classification. *Vision Research, 34*, 669–687.

Sagi, D. & Karni, A. (1993). The time course of learning a visual skill. *Nature, 365*, 250–252.

Shiu, L. P. & Pashler, H. (1992). Improvement in line orientation discrimination is retinally local but dependent on cognitive set. *Perception & Psychophysics, 52*, 582–588.

Taylor, M. M. & Creelman, C. D. (1967). PEST: Efficient estimates on probability functions. *Journal of the Acoustical Society of America, 41*, 782–787.

Weiss, Y., Edelman, S. & Fahle, M. (1993). Models of perceptual learning in vernier hyperacuity. *Neural Computation, 5*, 695–718.

Wilson, H. R. (1986). Responses of spatial mechanisms can explain hyperacuity. *Vision Research, 26*, 453–469.

Acknowledgements—We thank Mrs H. Weller, Mrs A. Hildinger and Mrs U. Weber for help during the experiments; Mrs H. Weller for artistic and secretarial help; Mr M. Repnow for writing the computer programmes that produced the stimuli; and Dr J. Harris for suggesting the experiments on the transfer of learning between offsets to the left and offsets to the right. Supported by the Deutsche Forschungsgemeinschaft (Fa 119/4-2, Zr 1/9-1) and the von Humboldt and Max-Planck-Society (Max-Planck Forschungspreis to M. Fahle and T. Poggio).